



**Europäisches
Patentamt**

**European
Patent Office**

**Office européen
des brevets**

Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03003692.5

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Anmeldung Nr:
Application no.: 03003692.5
Demande no:

Anmeldetag:
Date of filing: 18.02.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Tropf, Hermann, Dr.-Ing.
Blumenstrasse 5
68789 St.Leon-Rot
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se référer à la description.)

Database and method for organizing data elements

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

G06F17/30

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT SE SI SK TR LI

EPO - Munich
41
18. Feb. 2003

5 Dr.-Ing. Hermann TROPF
February 18, 2003
T50310EP

10 **Database and method for organizing data elements**

The present invention relates to a database system and a method for organizing data elements according to a Hilbert curve.

15 Organizing data elements in an efficient way is crucial for large databases in order to minimize data processing time. In particular, the number of accesses for retrieving data elements in a database is to be minimized when searching. The organization of data in a single dimension in a database is one way to achieve shorter access times. For mapping multidimensional data to one dimension, Z-ordering has been introduced in the last few years into commercial databases. It is supported by intuition and by theoretical studies that Hilbert ordering would be best, but the known efficient solution (1981 by the applicant) to the key problem of finding, from a given point, the next value in a multidimensional query rectangle has been reported to be complicated to be applied to Hilbert indexing.

25 Prior art Databases for multidimensional access may be roughly categorized into two different approaches:

A) Mapping multi dimensions onto one dimension using a space filling curve and sorting the records according to the one dimensional index. H. Tropf, H. Herzog: "Multidimensional Range Search in Dynamically Balanced Trees", *Agewandte Informatik* 2/1981, pp. 71-77 (in the following referenced by [1]), J. A. Orenstein: "Spatial Query Processing in an Object Oriented Database System", *ACM SIGMOD Int. Conf. on Management of Data*. 1986. pp. 326-336 (in the following referenced by [2]), and DE 196 35 429 (in the following referenced by [3]) with subsequent V. Markl: MISTRAL: "Processing Relational Queries using a Multidimensional Access Technique", Doctoral thesis, Techn. Univ. Munich, Germany. Infix Verlag, Sankt Augustin, Germany (in the following referenced by [4]), and V. Markl, F. Ramsak:

5 "Universalschlüssel: Datenbanksysteme in mehreren Dimensionen", c't 1/2001 pp. 174-179.
Heise Verlag, Hannover, Germany (in the following referenced by [5]) rely on Z-ordering (also
called Morton ordering) which is simply realized by bitwise interleaving the keys. (in the
following referenced by [6]). C.Faloutsos: "Multiattribute Hashing Using Gray Codes", Proc. of
the ACM SIGMOD 1986 Int. Conference on Management of Data.pp 227-238. (in the following
10 referenced by [6]) works on Gray coded blocks of bit interleaved data. J. K. Lawder, P. J. K.
King: "Querying Multidimensional Data Indexed Using the Hilbert Space filling Curve",
SIGMOD Record vol. 03 No. 1, 2001, pp. 19-24 (in the following referenced by [7]) and US
patent application 2003/0004938 A1 (in the following referenced by [8]) works on Hilbert
ordered data.

15

B) Many special structures for multidimensional data have been devised, most of them are
descendants of Kd-trees or R-trees, see <http://www.comp.nus.edu.sg/~liaojiro/tree.htm> (in the
following referenced by [9]) for a bibliography of 36 different tree types, H. Samet: "The
Design and Analysis of Spatial Data Structures", Addison-Wesley 1989 (in the following
20 referenced by [10]) or J.V. Gaede, O. Guenther: "Multidimensional Access Methods", ACM
Computing Surveys, Vol. 30 No. 2, June 1998. pp. 170-231. (in the following referenced by
[11]) for discussion.

The big advantage of A) over B) is that any tree balancing mechanism can be used to efficiently
25 handling dynamic data, hence also B-type trees which are widely used in commercial databases.
This is due to the fact that the mapping is independent of onedimensional data structuring.

Using Z-ordering with search trees has first been proposed in [1] (called "bit interlacing"); Fig.
1 shows the recursive Z-form of Z-ordered data for a 2D example. Z-ordering is one of the few
30 spatial access methods that has found its way into commercial database products ([11] section
4.3; R. Pieringer, K. Elhardt, F. Ramsak, V. Markl, R. Fenk, R. Bayer: Transbase: "A Leading-
edge ROLAP Engine Supporting Multidimensional Indexing and Hierarchy Clustering",
announced for: 10. GI Fachtagung Datenbanksysteme fuer Business, Technologie und Web.
26.-28.2.2003 (in the following referenced by [12]), for example now in use by the e-plus
35 mobile communication network in order to dynamically evaluate connections after geographical

5 and other criteria Transaction Software GmbH, Munich, Germany, www.transaction.de; www.transbase.de (in the following referenced by [13]). Insertion, deletion and exact search are done as usual, with logarithmic complexity. Range searches with small hypercube ranges have experimentally shown to be done in logarithmic expected time [1].

10 It seems clear from intuition and has been supported by theoretical studies that Hilbert ordering would be best (see, e.g. citations in [7]); Fig. 2 shows the recursive U-form of Hilbert ordered data for a 2D example. However, the efficient [1] solution to the key problem of finding, from a given point found, F, the next one in a multidimensional query rectangle, has been suspected [4] and reported [7] to be complicated to be applied to Hilbert ordering.

15 Therefore, it is an object of the present invention to provide a database system and a method for organizing indices of data elements according to a Hilbert curve which allows shorter access times to data elements being stored in the database system.

20 This object is achieved by a database system and method according to the independent claims. Further embodiments are defined in the dependent claims.

According to the invention, a database system is provided for organizing data elements according to a Hilbert curve, said data elements being representable by a plurality of

25 components, said database system comprising:

first means for generating a plurality of bitblocks by bitwise interleaving the components of the data elements;

second means for applying a fliprot transformation to a first bitblock;

said fliprot transformation comprising a flip transformation and a rot transformation, said flip

30 transformation indicating which bits of said bitblock are to be inverted, said rot transformation indicating which bits of said bitblock are to be interchanged;

third means for obtaining, for each further bitblock, a fliprot transformation by a concatenation of two or more fliprot transformations; and

fourth means for applying each fliprot transformation to its corresponding bitblock;

5 whereby the bitblock bits determine the organization of said data elements according to said Hilbert curve.

In a special aspect, said rot transformation indicates cyclically shifting the bits of said bitblock.

10 In a further special aspect, organizing is at least one of searching, sorting, storing, retrieving, inserting, deleting, querying, range querying, data elements in said database system.

Yet further, said organization comprises sorting said data elements into a binary tree or into a B-type tree.

15

According to the invention, the method of organizing data elements of a database according to a Hilbert curve, said data elements being representable by a plurality of components, comprises the following steps:

generating a plurality of bitblocks by bitwise interleaving the components of the data elements;

20 whereby a predetermined fliprot transformation is applied to a first bitblock;

said fliprot transformation comprising a flip transformation and a rot transformation, said flip transformation indicating which bits of said bitblock are to be inverted, said rot transformation indicating which bits of said bitblock are to be interchanged;

25 for each further bitblock, a fliprot transformation is obtained by a concatenation of two or more fliprot transformations;

and each fliprot transformation is applied to its corresponding bitblock;

whereby the bitblock bits determine the organization of said data elements according to said Hilbert curve.

30 The invention comprises also computer-readable data storage medium for storing program code for executing, when being loaded into a computer, according to the inventive method.

Thus, an algorithm is presented to solve this key problem for Hilbert ordering with advantage over the Lawder [7,8] algorithm. It is linear with the number of dimensions and linear with the
35 coordinate's values wordlength. The method is generic in the sense that it can be switched from

5 Hilbert indexing to alternative hierarchical indexing schemes by just changing two statements in the kernel algorithm (Z-indexing being the simplest case). As a side-product, we have found a tiny algorithm for calculating the n-dimensional Hilbert index. First experimental results are available for multidimensional range searching.

10 The invention is described in more detail with reference to the drawings, wherein:

Fig. 1 shows an example of Z-indexing in 2D with values, $x=0..7$, $y=0..7$;

Fig. 2 shows an example of Hilbert indexing in 2D with values, $x=0..7$, $y=0..7$;

15 Fig. 3 shows a 3D Hilbert Cube with 2 bit resolution;

Fig. 4 displays Gray Code examples;

20 Fig. 5 displays an example for a flip rot transform in 3D; and

Fig.6 serves for explanation for calculating the standard solution.

BIGMIN: next index point in query rectangle

25 In the following, the key problem mentioned above is discussed; we refer to it as BIGMIN problem in the following.

Range search is important when processing multidimensional point data; it is not only directly important, it serves also as basis for doing nearest neighbour or similarity searches efficiently.

30 Regardless what data balancing mechanism is used (binary, B-type or other), regardless what indexing scheme is used, multidimensional range searching ends up in the problem of efficiently finding, from a given point F found so far, the next one (after indexing scheme) which is in a

5 multidimensional query rectangle. Stated otherwise, it is the rectangle point with minimum index bigger than the index of F; it is called BIGMIN in the following.

When searching is done in left-right manner, BIGMIN is the only thing needed. If searching is done top-down, as usual with search trees, it is helpful to calculate also the opposite thing, the
10 rectangle point with maximum index, smaller than the index of F. This point is called LITMAX in the following.

At first sight, bit interleaving seems to have substantial difficulties when the query range overlaps the "strong" borderlines with large Z-value value jumps. On basis of the
15 BIGMIN/LITMAX calculation, much of the search tree can be pruned: Suppose that a node F with Z-value 19 has been found; then BIGMIN is 36, LITMAX is 15. To the left of F, only Z-values 12..15, to the right only values 36..39 must be searched. Performing an efficient BIGMIN / LITMAX calculation is therefore a key problem to range searching.

20 Changing the [1] basic tree search algorithm slightly to comply with our Hilbert indexing requirements (dealing with points instead of indexes), range search with BIGMIN/LITMAX is briefly stated in pseudo code as follows: (Plo / Phi is the point in the rectangle with lowest / highest Hilbert value in the rectangle), H(P) is the Hilbert index of a point P):

25 Algorithm 1:

calculate Plo, Phi

Range (P, Plo, Phi):

case 1: $H(P) < H(Plo)$:

30 Range (High Son of P, Plo, Phi).

case 2: $H(P) > H(Phi)$:

Range (Low Son of P, Plo, Phi).

case 3: $H(Plo) \leq H(P) \leq H(Phi)$:

report P if it lies in the query hyper rectangle

35 Compute BIGMIN and LITMAX

5 Range (Low Son of P, Plo, LITMAX)
 Range (High Son of P, BIGMIN, Phi)

10 BIGMIN range searching with B-type trees:

The modification of the algorithm for B-type trees (developed for external searching), where each node has more than one record, is obvious. It is shown by means of the following typical situations:

15 Situation a): Nodes have up to 1 son per node record. A node P has records R_i with H-value $H(R_i)$. A Record R_i has up to 1 son S_i . Any H-value in the subtree of R_i are between $H(R_{i-1})$ and $H(R_i)$. (This corresponds roughly to the definition of a B-tree, neglecting the an additional rightmost son in order to make the description more readable; B*-trees are essentially the same
20 but with a different minimum filling degree).

Algorithm 2a is as follows:

 calculate Plo, Phi
25 Range (P, Plo, Phi):
 for each R_i in P do
 { Report R_i if it lies in the query hyper rectangle.
 if $H(Plo) < H(R_i)$ and $H(Phi) > H(R_{i-1})$ then
 {
30 compute BIGMIN with $H(R_i)$
 Range(S_i , BIGMIN, Phi)
 }
 }

5 The application to B+ trees (data stored in the leaves; pointers provided to the subsequent leave)
is along the same lines.

Situation b): A node P has up to 2 subtrees. H-values in P are greater than any H-values in any
nodes of the left subtree; H-values in P are smaller than any H-values in any nodes of the right
10 subtree. The lowest H-value in P is called Hmin(P), the highest H-value in P is called Hmax(P).

Algorithm 2b is as follows:

```
calculate Plo, Phi
15 Range (P,Plo, Phi):
  case 1: Hmax(P)<Plo:
    Range (High Son of P, Plo, Phi).
  case 2: Hmin(P)>HMAX
    Range (Low Son of P, Plo, Phi).
20 case 3: Plo<=H-Value(P)<=Phi
    Report all records in P that lie in the query hyper rectangle.
    Compute BIGMIN with Hmax(P)
    Compute LITMIN with Hmin(P)
    Range (Low Son of P, Plo, LITMAX)
25 Range (High Son of P, BIGMIN, Phi)
```

BIGMIN solution for Z-indexing

Now, the 1981 solution for Z-indexing [1] is recalled as much of its ideas can be applied to
Hilbert indexing. The [1] approach is recalled because the basic concepts are easier seen with Z-
30 coding; after that, the application to the more complicated Hilbert indexing will be described.

The calculation of BIGMIN for Z-indexing is realized as a binary search with stepwise bisecting
the data cube. Point F data (Z(F)) are bitwise scanned in interleaved order; at each step, the

5 position of F and of the query rectangle is examined in relation to the bisecting line. The
rectangle is given by its MIN/MAX Z-value corners. MIN,MAX data are also bitwise scanned.
There is a staircase borderline bisecting Z-values > Z(F) from Z-values < Z(F). F is given in
brackets in the examples following.

10 Six cases are possible when searching BIGMIN:

Case A: F is left of bisection line (Fbit=0).

Case A1: Range is totally left of bisection line (Fbit=0 MINbit=0 MAXbit=0) Example:

15 18 24 26 |
 19 25 27 |
 22 28 30 |
(21)

20 Everything is going on in the low section. Continue.

Case A2: Section Line crosses query range. (Fbit=0 MINbit=0 MAXbit=1) Search continues to
the left; but two cases possible, but not yet distinguishable:

A2a: The staircase crosses the query region straight, exactly along the section line. Example:

25 7 13 15 | 37 39
 18 24 26 | 48 50
 19 25 27 | 49 51
 (29)

30 If this is the case, BIGMIN is the lowest possible value in the high section (37). This value is
calculated by simply loading 1000.. into MIN, called "candidate", starting from the actual bit
position.

A2b: The staircase crosses the left query region in staircase form Example: (Here, BIGMIN will
35 be finally 24)

5

7	13	15		37	39
18	24	26		48	50
19	25	27		49	51
22	28	30		52	54

10 (23)

BIGMIN is in the left section. The rectangle is shrinked. MAX jumps from 54 to 30. This jump is simply done by loading 0111.. into MAX, starting from the actual bit position.

15 Case A3: Range is totally right of bisection line (Fbit=0 MINbit=1 MAXbit=1). Example:

(14)
37 39
48 50
49 51

20 (38)

MIN has become greater than Z(F). BIGMIN:=MIN. finish. Remark: This can happen due to shrinking the rectangle.

25

Case B: F is right of section line (Fbit=1)

Case B1: Range is totally left of bisection line (Fbit=1 MINbit=0 MAXbit=0). Example:

(38)
5 7 13 15
16 18 24 26
17 19 25 27
20 22 28 30

30

5 MAX has become lower than Z(F) (this can happen due to shrinking the rectangle). BIGMIN must have been saved before. Report BIGMIN as saved. Finish.

Case B2: Range is totally right of bisection line
(Fbit=1 MINbit=1 MAXbit=1) Example:

10

```
| 48    50
| 49    51
| 52    54
(53)
```

15

Everything is going on in the high section. Continue scanning.

Case B3: Section Line crosses query range. Fbit=1 MINbit=0 MAXbit=1

20 (42)

```
18    24    26 | 48    50
19    25    27 | 49    51
22    28    30 | 52    54
```

25 If this is the case, BIGMIN must be in the high section. continue searching in the high section. The rectangle is shrunked. MIN jumps from 18 to 48. This value is calculated by loading 1000.. into MIN, starting from the actual bit position.

The LITMAX computation is analogous, with symmetries. The complete BIGMIN/LITMAX
30 decision table can be found in [1]. Z-BIGMIN/LITMAX algorithm as recalled is linear with the number of dimensions and linear with the coordinate value's wordlength (supposed proper realisation of the LOAD function).

We will follow these guidelines for doing the same thing for Hilbert coded data.

35

5 Alternative representation of Hilbert indexing

In the following, we introduce an alternative view of Hilbert indexing that serves as basis for the algorithm described afterwards.

3D example

10 The method presented in the present application is based on a special representation of Hilbert indexing which is described below.

The Hilbert curve is a space filling curve (each data point visited exactly once) with

- only single steps in exactly one dimension
- hierarchically bisecting the data cube.

15

Let us first take a look on Fig. 3 for 3D with 2 bit resolution. We think the 3D, 2 bit resolution data cube as consisting of 8 subcubes with 1 bit resolution each. The Hilbert curve is a walk from one subcube to the next; the main bisection is between the front and the rear subcubes in the figure; in each side the subcubes are visited in a U-shaped manner. The subcubes themselves are visited internally in the same manner, mirrored and/or rotated the way as requested by their entry and exit position (in Fig. 3 only the internal curve of the first subcube is shown)

20

Gray codes with flip/rot representation

25 Turning to a bit oriented view, Hilbert indexing can be regarded as bit interleaved Gray Codes with special requirements on the Gray Codes used.

Gray coding means coding a sequence the way that at each step only one bit changes. For a cyclic Gray code, in addition, only 1 bit is different when comparing the first and the last code (Fig. 4, Example 1). A given cyclic Gray code can be doubled by adding one bit with first half and second half different, and mirroring the rest (Example 1

30

- -> Example 2).

5 Gray codes that allow a columnwise hierarchical decomposition of the indices without considering wrap around are called G-code in the following (examples 1..4, example 5 is a counterexample). The classic example 1 or 2 code is called (standard) Gray code in the following.

10 A G-code remains a G-code if a column is inverted. A G-code remains a G-code if any columns are exchanged (with rotations as a special case). Inverting one or more columns is done by XORing the corresponding bits with 1. An array of bits indicating which column of a code to be XORed, is called Flip in the following. The procedure is called flipping. Flipping example 2 by 101 yields Example 3.

15 The problem discussed in this application is solved by only considering rotations, we need not think about exchanging. When handling rotations, we only think of the no. of columns it has been rotated. We define left rotations positive (in the direction of more significant bits of standard Gray code). Rotating Example 3 by +1 yields example 4. To describe the example 4 G-
20 code, we simply write (101/+1) denoting that the standard Gray code has been flipped by 101 and then rotated by +1.

flip/rot representation for Hilbert cube

With Z-indexing, we strictly scanned the data bitwise in interleaved order, beginning with the most significant bit, e.g. for three dimensions: xyzxyzxyz.... We can look at it bitblockwise: xyz
25 xyz xyz; this is what we do to cope with Hilbert indexing. Each bitblock represents a one bit (sub)cube with one bit resolution.

Note that the decimal numbers given at the left in Fig. 4 are the indices, the codes are, in binary interpretation, the bit interleaved geometric coordinates.

30 For Hilbert indexing the Fig. 3 cube, we take the Gray code for the main bitblock denoting the sequence of subcubes. For each of these code values, we have to find a G-code the way it complies with the Hilbert indexing requirements.

5

The Hilbert requirements are explained now with reference to Fig. 5, wherein transformation (lmn/r) means flip with lmn , rotate by r $Tab(i)=(lmn/r)$ means that the transformation (table) for index i is (lmn/r) . The three Hilbert requirements are:

- 10 (1) Main entry and main exit are main cube corners, so there the coordinate values are extreme, i.e. either 000... or 111... Viewed bitblockwise that means that the bitblocks must be the identical (see positions (a)) in Fig. 5. (2) When changing from subcube to subcube, exactly one coordinate changes by a single geometrical step. Therefore the changing coordinate bit must do just the opposite thing of the main cube bit, see positions (b) in Fig. 5. (3) The codes need to be
15 cyclic. Therefore exactly one bit must change at positions (b) in Fig. 5. Two cases are possible: Due to the second requirement, exactly one of the last row bit is required to be different to the corresponding first row bit. If the bit under consideration happens to be different, the remaining bits must be copied from the first row. Otherwise we have free choice which of the remaining bits to make different.

20

Without loss of generality we assume that the first bitblock is a standard Gray code. A solution for the second bitblock column, under this assumption, is called a standard solution in the following. Once a standard solution for a given number of dimensions is known, i.e. the sequence of subcube coordinate transforms (mirroring and rotations) in the main cube, the
25 solution for any deeper subcube can be calculated directly by a concatenation of flip/rot transforms. This is shown in Fig. 5 for subcube indexed with 6. The concatenation is surprisingly easy and can be found in algorithm 1.

To make plausible that concatenation works: imagine for the moment that the 2nd bitblock code in question (Tab_6) would be the standard Gray code instead of (110, 010, 011, 111, ..). Then the 3rd bitblock G-code would be the Tab_3 -Standard G-code. Then imagine that both 2nd and 3rd bitblock G-codes are flipped and rotated by the Tab_6 flip/rot to fit the 2nd bitblock with the first bitblock (parallel flipping/ rotating does not change the relations between G-codes under consideration).

35

5 For simplicity of description assume first that for the given number of dimensions a standard
solution flip table / rotation table is given; in the following algorithms we provide flip tables and
rotation tables as constant arrays, for 2 or 3 dimensions. Later, we describe how flip and rot
standard solution values are calculated "online" without the aid of precompiled tables
(Calculating the standard solution).

10

Calculating the Hilbert index

An algorithm that follows the above concepts is given as Algorithm 1.

Calculating the Hilbert index is not really needed for the problem discussed, but this algorithm
15 serves as framework mechanism for the following algorithms to plug in specific blocks at
places. Wordlength considerations are only critical when really calculating the Hilbert value.

The following algorithms are given in plain Pascal, with Shift and AND/OR/XOR operations
allowed as in Borland Pascal. Throughout these algorithms, hi/lo refers to Hilbert index,
20 right/left to coordinates. Local comments are given within the source code, general comments at
the end of the source code.

Global declarations are as follows, (auxiliary functions and tables to be found in detail in the
appendix):

25

Algorithm 1:

Type/Const/Variable Declarations:

30 (*constants to choose:*)
const ndims = 3; (*no. of dimensions*)

(*beware wordlength for calc_H; longint used here as max. wordlength.*)
const bitresolution= (sizeof(longint) * 8) div ndims;

```

5  const initialbit= bitresolution -1;

(*types to work with:*)
type point = array [1..ndims] of word;
type rectangle = record
10     left, right: array [1..ndims] of word;
    (*left = low border, right= high border,
       viewed in point coordinates*)
    end; (*record*)

15  (*dependant:*)
type block    = array [1..ndims] of boolean;
const G_CodeLength = 1 shl ndims;

Semantics of the calc_H variables:

20  bitpos: integer running bit position within word length
    d: running. dimension for Gray converted data
    drot: from d back rotated index for original data access.
        drot[d] is where the flip for d was active

25  G_index: Gray-code Index for next block
    Flip, rot: integer; running Hilbert index representation
        derived from old one and from G-index
    toggle: does Gray coding: going hi half, Gray
        codes mirror in the following. toggle inverts each
30     time when going high half.
    data: input data converted to Gray representation.
    indblk: block; array of hi/lo decisions in d order.
    inverted: boolean; tells if hi/lo inverted against right/left
    mask: to fetch the bits at bitpos.

35  result: beware wordlength! Only needed if Hilbert index

```

5 is really calculated, not needed for further algorithms.

Inverted: indicates if left/right means "hi/lo" or "lo/hi".

Inverted and drot are only needed for working blocks of further algorithms.

10 This is the function for calculating the n dimensional Hilbert index (algorithm 1):

```
function calc_H(p: point): longint;
(*calculates Hilbert index for point data*)
var bitpos:integer;d: integer;G_index: word;
15 flip: block; rot: integer; toggle: boolean;
data: block;indblock: block;
drot: array [1..ndims] of integer;inverted: boolean;mask: word; result: longint; h: integer;
begin
    result:=0; for h:=1 to ndims do flip[h]:=false; rot:=0;
20
    for bitpos:=initialbit downto 0 do
    begin
        mask:=1 shl bitpos;
        for d:=ndims downto 1 do
25
            (*this is the generate data block:*)
            data[d]:=(p[d] and mask)<>0;
            fliprot((*var*) data, flip, rot);

30 toggle:=false;
        for d:=ndims downto 1 do
        begin
            (*only for other procedures:*)
            drot[d]:=mod_( (d-1 - rot), ndims) +1;
35 inverted:= flip[drot[d]] XOR toggle;
```

```

5      (*here optionally comes the working block*)

      indblk[d]:=data[d] XOR toggle; (*true if go hi*)
      toggle:=toggle XOR data[d]; (*toggle for hi data*)

10     (*This is the update result block*)
      if indblk[d] then
          result:=result or (1 shl ((bitpos*ndims)+d-1) );
          calc_H:=result;

15     end; (*for d*)

      g_index:=0;
      for d:=ndims downto 1 do
20     if indblk[d]
          then g_index:=g_index or (1 shl (d-1));

          for d:=ndims downto 1 do
              if indblk[d] then g_index:=g_index or (1 shl (d-1));

25     concat(flip,rot, fliptab[G_index],rottab[G_index],
              (*var*) flip, (*var*) rot);(see appendix)

          end; (*for bitpos*)
30     end; (*calc_H*)

```

BIGMIN/LITMAX solution for Hilbert indexing

35

5 Now, the algorithms are described in detail, based on the concepts introduced above (BIGMIN/LITMAX solution for Z-indexing, Hilbert indexing, resp.). The description thus far relies on a table, precompiled once for a given number of dimensions;

What we need is two things:

10

(Problem 1) A function that tells which of two data points has the greater Hilbert value. This is needed for inserting, deleting and exact searching (the Hilbert value itself is not really needed), and (Problem 2) an efficient H-BIGMIN (and H-LITMAX) computation.

comparing data points after Hilbert index

15 Problem 1 is solved with

```
function greater(p, p1: point): boolean;  
(*true if hilbert value of p > Hilbert value of p1*)
```

20 which is easily accomplished by the following replacements in the calc_h function:

```
(*this is the generate data block:*)  
for d:=ndims downto 1 do  
begin  
25   p_in_right[d] :=(p [d] and mask) <> 0;  
   p1_in_right[d]:=(p1[d] and mask) <> 0;  
end;  
data:=p_in_right; fliprot((*var*) data, flip, rot);
```

```
30 (*This is the update result block:*)  
if p_in_right[drot[d]]<>p1_in_right[drot[d]] then  
begin  
   greater:=p_in_right [drot[d]] XOR inverted; exit;  
end;
```

5 greater:=false;

BIGMIN/LITMAX algorithm for Hilbert indexing

In order to solve problem 2, we first solve Problem 2a: - within a rectangle, find the coordinates with lowest Hilbert index. Based on its solution we develop a solution of the problem 2b, the

10 BIGMIN problem:

searching rectangle point with minimum index

The lowest index in a query rectangle is no more simply the low rectangle corner index (as is the case with Z-indexing). A rectangle is represented by its outer borderline coordinates (left and right for each dimension). The following algorithm calculates the coordinates of the point

15 with lowest Hilbert value within a rectangle (problem 2a).

It is basically the bitwise scanning of algorithm 1, with a binary search cutting the rectangle at each step if it overlaps the bisecting dimension.

20 The auxiliary functions forcehigh and forcelow are similar to the Load functions used for Z-indexing, but they will be applied to rectangle data instead to point data.

The procedure calc_lowest_Hpoint_in_rectangle needs a few additional local variables:

25 procedure calc_lowest_Hpoint_in_rectangle(r: rectangle; var H_point: point);
 var (*declarations see calc_H; in addition:*)
 in_right, in_left: bitblock; in_lo, in_hi: boolean;
 data_in_right, data_in_left: bitblock;

30 The generate data block is:

for d:=ndims downto 1 do
 begin

```

5      in_right[d]:= (r.right [d] and mask) <> 0;
      in_left [d]:= (r.left [d] and mask) = 0;
      end;

      (*generate normalized data to Gray*)
10     data_in_right:=in_right; fliprot((*var*) data_in_right, flip, rot);
      data_in_left :=in_left; fliprot((*var*) data_in_left, flip, rot);

```

The working block is

```

15  if NOT inverted
      then begin
          data[d]:=data_in_right[d];
          if in_right[drot[d]] and in_left[drot[d]] then
              forcei((*var*) r.left[drot[d]], bitpos);
20      end
      else if in_right[drot[d]] and in_left[drot[d]]
          then begin
              data[d]:=not data_in_left[d];
              forcele((*var*) r.right[drot[d]], bitpos);
25      end
          else if in_right[drot[d]] and (not in_left[drot[d]])
              then data[d]:=data_in_right[d]
              else if (not in_right[drot[d]]) and in_left[drot[d]]
                  then data[d]:=not data_in_left[d]
30      else error ('1');

```

The update result block is:

```

if indblk[d] XOR inverted (*convert bitset decision back*)
35  then H_point[drot[d]]:=H_point[drot[d]] or mask;

```

5

BIGMIN/LITMAX algorithm

The BIGMIN problem 2b is then solved as follows:The same basic idea as with Z-indexing described above, using the hilbert calculation mechanisms of the foregoing algorithms, but there is a serious complication: The candidate point becomes more difficult to calculate.

10

It is not good to calculate the candidate point immediately when a candidate must be saved: maybe it is not needed at all, maybe there will come better candidates while searching. If we would do the calculation immediately, the procedure would become quadratic with the number of dimensions.

15

We do now the following: When a candidate must be saved, we simply save it in form of the sub-rectangle in which it is the lowest/highest value. If it turns out that this candidate is the solution, it is still the right time to do the calculation. If a better candidate shows up, we simply overwrite the candidate's rectangle data. So in the end at most one candidate must be calculated, and the procedure becomes linear.

20

In the heart of the algorithm, again the 6 cases are distinguished as with Z-indexing explained above.

25 Here we give only one example for space reasons:

F is in Low section, rectangle overlaps sections

(18)

30

0	3	4	5		58	59	60	63
1	2	7	6		57	56	61	62
14	13	8	9		54	55	50	49
15	(12	11	10		53	52)	51	48
16	(17	30	31		32	33)	46	47
19	18	29	28		35	34	45	44

35


```

5      20  23  24  27 | 36  39  40  43
      21  22  25  26 | 37  38  41  42

```

BIGMIN is either in low section or the MIN of rectangle in hi section. Save candidate in high, cut hi section from rectangle, go lo section.

10

The BIGMIN procedure:

```

procedure calc_BIGMIN(r: rectangle; F: point; var BIGMIN: point);

```

(*F is the point found in a search tree. precondition:

15 H-index of F is between highest/lowest H-indices in r, but geometrically not in r*)

needs a few local variables in addition to the procedure calc_lowest_Hpoint_in_rectangle:

```

var F_in_right: bitblock; F_in_hi: boolean; cand: rectangle;

```

20

The generate data block is:

```

for d:=ndims downto 1 do
  begin in_right [d]:= (r.right [d] and mask) <> 0;
25      in_left  [d]:= (r.left  [d] and mask) = 0;
      F_in_right[d]:= (F      [d] and mask) <> 0;
  end;
data:=F_in_right; fliprot((*var*) data, flip, rot);

```

30 the working block, including result calculation is

if NOT inverted

```

  then begin F_in_hi:=F_in_right [drot[d]];

```

```

        in_hi :=in_right  [drot[d]];

```

35

```

        in_lo :=in_left   [drot[d]];

```

```

5      end
      else begin F_in_hi:=not F_in_right [drot[d]];
                in_hi  :=in_left      [drot[d]];
                in_lo  :=in_right     [drot[d]];
                end;
10  if F_in_hi then (*implies go hi*)
      begin
        if not in_hi then(*search fails; rep. Min in cand.*)
          begin
            calc_loHpoint_in_rect(cand, (*var*) BIGMIN); exit;
15      end;
            if in_lo then
              if not inverted
                then forceri((*var*) r.left [drot[d]], bitpos)
                 else forcele ((*var*) r.right[drot[d]], bitpos);
20      end (*di_in_hi*)
            else
              begin (*di_in_lo*)
                if not in_lo then(*search fails;cand.is BIGMIN*)
                  begin calc_loHpoint_in_rect(r, (*var*) BIGMIN);
25                  exit;
                  end;
                if in_right[drot[d]] and in_left[drot[d]] then
                  begin cand:=r; (*save candidate hi, cut lo, go lo:*)
                    if not inverted then
30      begin forcele((*var*) r.right [drot[d]], bitpos);
                          forceri((*var*) cand.left [drot[d]], bitpos);
                    end else
                      begin forceri((*var*) r.left [drot[d]], bitpos);
                                forcele((*var*) cand.right[drot[d]], bitpos);
35      end;

```

5 end;
 end; (*di_in_lo*)

The H-LITMAX computation is the same thing with inverted thinking.

10 Remarks:

A number of technical improvements are possible: When rotating the data, copying can be omitted by merely rotating working indices.

15 Another technical improvement is that the candidate calculation is only necessary to be done starting with the bit position at which the candidate has been created. To do this, the bitposition and the running flip/rot state have to be saved together with the candidate. BIGMIN and LITMAX can be calculated in parallel because F is the guiding point. We only need two candidates of course, one for BIGMIN and one for LITMAX, exit to candidate calculation to be
20 coordinated with bookkeeping.

We did not present the algorithm with those technical improvements in order to make the description better understandable.

25 Calculating the standard solution

Now, we describe how this precompiling can be circumvented by replacing the table lookup by a function call that is free from iteration or recursion.

5 When precompiling, Lawder [7,8] uses state transition diagrams that are much more complicated than the data described here, derived from so-called generator tables. The Fig. 5 column 2 data are part of what Lawder calls a generator table. Lawder has observed that there is a system within this column. Influenced by his work at this point, we present a somewhat different view of the same sequence. Based on this view, we provide an algorithm for
10 calculating the G-code representation just for a given index, without calculating the table column as a whole thing.

We start with a primitive cell Fig. 6(a) which is the representation of a 1D, 2-bit data cube conforming with the Hilbert indexing requirements. Going to 2D, at first the whole thing is
15 mirrored and a 00..0,11..1 sequence added (b) (as with Gray coding). Then, in order to comply with the Hilbert indexing requirements, we invert the outermost bits (">", "<") at the mirror point, see (c). The 3D standard solution is shown in (d).

The algorithmic solution for a given index i is as follows: Entry code and exit code are set to the
20 gray code of index i . Apart from LSB bits, the Entry code bits are inverted at places where the binary representations of $i-1$ to i changes from 0 to 1. The exit code bits are inverted accordingly for a binary 0-1-change from i to $i+1$. If such a bit is inverted, then also its LSB is inverted. Then, flip is simply the entry code value and rot is the place where entry and exit bits differ. We do not give the source code due to lack of space and because transformation to source code is
25 easy.

Why it works:

- the first bit combination is always 0000...
- the last bit combination is always 1000...
- 30 - mirroring does not change the conditions apart from the mirror point.
- At the mirroring point, after mirroring, inverting the leftmost bit makes it comply with Hilbert requirement 2 (see above).
- 35 - This, however, introduces an inversion to the

- 5 foregoing bit which is compensated by undoing the
inversion that always takes place at the rightmost bit
of the mirror point, as it stems from the primitive
element.
- 10 Note that there are many solutions that comply with the Hilbert indexing requirements (2D has
one solution, 3D has 2 solutions and one dead end when doing exhaustive backtracking); we
consider the solution presented here as being canonical as it is a minimal amendment to Gray
coding.
- 15 To avoid precompiling the flip/rot table for a given dimension, just replace the fliptab/rotab
table lookup in algorithm 1 by corresponding flipfunc/rotfunc function calls.

First experimental results

- Performance is measured in terms of the number of nodes inspected. Test data are generated
20 with a pseudo random generator for both the data in the database and the query data, both over
the whole range of the data cube.

- First experimental data give a mean 10 % improvement of Hilbert ordering over Z-ordering
(single cases are possible where Hilbert ordering is even worse than Z-ordering). Our
25 experiments are done with up to 10 dimensions.

A further application of the inventive concept is to nearest neighbour searching,.

Discussion; related work

- 30 To do range searching in Z-data, [2] decomposes the query hyper rectangle into a sequence of
elements each with consecutive Z-values, in order to do an optimized merging of the sequence
with the sorted data. The sequence can be very large as there are many (possibly very small) Z-
value holes in the rectangle. We did not consider it for Hilbert indexing as even for Z-indexing

5 it compares unfavorably with the earlier BIGMIN calculation approach. When used with B*-trees, Bayer terms "UB-trees" (Universal B-trees). Bayer [3] proposed a procedure to do the job which turns out to be exponential with the number of dimensions ([4] p. 123). In course of the European MISTRAL project <http://mistral.in.tum.de>, this "critical part of the algorithm" has been replaced by a linear "bit oriented algorithm" not described in detail ([4] p. 124).

10

Lawder uses either precompiled state transition diagrams, or he does a direct calculation that needs iteration. The method described here differs basically from the Lawder approach:

15 We do by means of a flip/rot representation and its very simple concatenation transform. We presented a fast noniterative calculation by means of a simple concatenation of a flip/rot representation so that precompiling does not make much sense; if precompiling is done anyway, the flip/rot tables are much compacter than state diagrams. Calculation is done for a given index, without calculating the table column as a whole thing.

20 We use the [1] candidate technique; what Lawder does using explicit backtracking is done by simply saving a rectangle's data as candidate. Another thing may be worth to be mentioned: when bisecting the space, Lawder uses two limits explicitly: max_lower and min_higher. We show by our algorithm that these limits are not really needed.

25 As our flip/rot transformation is free from recursion or iteration, the whole BIGMIN/LITMAX algorithm becomes linear with the number of dimensions and linear with the coordinate's values wordlength; this is true although working bitwise, as in a technically optimized version rotations are done in one step by changing the working indices accordingly, without copying data.

30 Last not least we do not necessarily process the search tree left-right; starting at the root of the search tree and working recursively to both sides with both BIGMIN and LITMAX is more convenient as skipping subtrees is done in a natural way.

35 The Lawder approach has been presented for B-type trees searching for the page key of BIGMIN's bucket (the page key is the key with minimum index within the bucket). We strictly

5 separate the search procedure from the BIGMIN calculation, thereby making considerations and
adaption to alternative data handling systems easier; we have shown how the concept applies to
both binary and B-type trees.

10 We have presented the method in a modular way; our algorithm is generic in the sense that it is
simple to be changed to alternative hierarchical indexing schemes by changing the fliprot and/or
toggle and/or concat lines of frame algorithm 1 suitably (for Z-indexing just cancel the latter
two lines).

15 As a side product, we have found a tiny algorithm for bitwise calculating the n-dimensional
Hilbert index.

Technical remarks:

20 We did not consider scaling. In real applications, scaling should be done the way that the data
cover the data cube nearly equally in all dimensions.

25 Both Z-indexing and Hilbert indexing apply also to negative and to real valued data. The only
thing that is requested is that the bits are accessed in the order of significance (start with the
exponent, MSB first, followed by the mantissa; invert sign bits).

30 For both Z-indexing and Hilbert indexing, bit interleaving is not done explicitly. We keep the
data as usual and just scan the bits in interleaved order. Z-indices or Hilbert indices are not
calculated explicitly, so there are no wordlength problems. BIGMIN and LITMAX values are
working records represented just as normal records.

35 In a technically optimized version just some additions and XORs do per resolution bit should
do. With the solution presented in this application it looks clear that for external storage the
overhead against Z-indexing pays, because it virtually vanishes in relation to the time needed
for disk accesses (question posed by [4] p. 190).

5 Appendix: auxiliary functions + tables

```
function mod_(a: integer; modulo: integer): integer;
(*modulo correctly for neg. values*)
begin
10   a:=a mod modulo; if a<0 then mod_:=a + modulo
                                else mod_:=a;
end; (*mod_*)

procedure rotateblock(var B: bitblock; r: integer);
15 (*rotates B by r*)
var hB: bitblock; h: integer;
begin
    hB:=B; (*copy: see text for technical improve-
            ments*)
20   for h:=1 to ndims do B[h]:=
        hB[mod_(h-1-r,ndims)+1];
        (*shift right is fetch left*)
end; (*rotateblock*)

25 procedure fliprot(var B: bitblock; flip: bitblock; rot: integer);
(*flips B with flip and then rotates by r*)
var h: integer;
begin
    for h:=1 to ndims do B[h]:=B[h] XOR flip[h];
30   rotateblock((*var*) B, rot);
end; (*fliprot*)

(*Example tables precompiled once for a given no. of dimensions, ndims = 2, 3:
function replacement see text*)
35 (*2D:*)
```



```

5  const fliptab: array [0..G_CodeLength-1] of
      bitblock =((false,false), (false,false),
                  (false,false), (true ,true ));
const rottab: array [0..G_CodeLength-1] of integer =(1, 0, 0, 1 );

10
(*or, 3D:*)
const fliptab: array [0..1 shl ndims-1] of bitblock =
((false,false,false),(false,false,false),
 (false,false,false),(true ,true , false),
15 (true ,true , false),(false,true ,true ),
 (false,true ,true ),(true ,false,true ));
const rottab: array [0..1 shl ndims-1] of integer =
( +2, +1, +1, 0,
 0, +1, +1, +2 );

20
procedure concat_fliprot(f1: bitblock; r1: integer;
      f2: bitblock; r2: integer;
      var f: bitblock; var r: integer
      );
25 (*concat flip/rot transforms f1/r1 and f2/r2 to single transform f/r:*)
(*f2 shifted back by r1, then f1 XOR f2. r=r1+r2. Result f order dependent!*)
var h: integer;
begin
    rotateblock((*var*)f2, -r1);
30  for h:=1 to ndims do f[h]:=f2[h] XOR f1[h];
    r:= mod_(r1+r2, ndims);
end; (*concat_fliprot*)

procedure forceri(var b: word; bitpos: integer);
35 (*"force right": forces highest possible value into b,

```

```

5   beginning with bitposition bitpos; bitpos = 0...*)
var mask: word;
begin
  if bitpos>(sizeof(mask)*8-1)then error('wordlength');
  mask:=1 shl bitpos;
10  (*force 1 into actual bitposition, e.g. .OR 001000..*)
  (*force 0 into the rest,          e.g. .AND111000..*)
  b:= b OR ( mask );
  b:= b AND (NOT (mask-1));
end; (*forceri*)

15
procedure forcele(var b: word; bitpos: integer);
(*"force left": forces lowest possible value into b,
  beginning with bitposition bitpos; bitpos = 0...*)
var mask: word;
20 begin
  if bitpos>(sizeof(mask)*8-1)then error('wordlength');
  mask:=1 shl bitpos;
  (*force 0 into actual bitposition, e.g. AND110111..*)
  (*force 1 into the rest,          e.g. .. OR000111..*)
25  b := b AND (NOT mask );
  b := b OR ( (mask-1));
end; (*forcele*)

```

5 Dr.-Ing. Hermann TROPF
February 18, 2003
T50310EP

Claims

10

1. A database system for organizing data elements (P) according to a Hilbert curve (H), said data elements (P) being representable by a plurality of components, said database system comprising:

15

first means for generating a plurality of bitblocks (xyz) by bitwise interleaving the components of the data elements (P);

second means for applying a fliprot transformation to a first bitblock;

20

said fliprot transformation comprising a flip transformation and a rot transformation, said flip transformation indicating which bits of said bitblock are to be inverted, said rot transformation indicating which bits of said bitblock are to be interchanged;

third means for obtaining, for each further bitblock, a fliprot transformation by a concatenation of two or more fliprot transformations; and

25

fourth means for applying fliprot transformations to their corresponding bitblock; whereby the bitblock bits determine the organization of said data elements (P) according to said Hilbert curve (H).

30

2. The database system according to claim 1, wherein said rot transformation indicates cyclically shifting the bits of said bitblock.

3. The database system of claim 1 or 2, wherein organizing is at least one of searching, sorting, storing, retrieving, inserting, deleting, querying, range querying, data elements in said database system.

35

- 5 4. The database system according to one of the preceding claims, wherein said
organization comprises sorting said data elements (P) into a binary tree or into a
B-type tree.
- 10 5. A method of organizing data elements (P) of a database according to a Hilbert
curve (H), said data elements (P) being representable by a plurality of
components, said method comprising the following steps:
- generating a plurality of bitblocks (xyz) by bitwise interleaving the components of
the data elements (P);
- 15 whereby a predetermined fliprot transformation is applied to a first bitblock;
said fliprot transformation comprising a flip transformation and a rot
transformation, said flip transformation indicating which bits of said bitblock are
to be inverted, said rot transformation indicating which bits of said bitblock are to
be interchanged;
- 20 for each further bitblock, a fliprot transformation is obtained by a concatenation of
two or more fliprot transformations;
and fliprot transformations are applied to their corresponding bitblock;
whereby the bitblock bits determine the organization of said data elements (P)
according to said Hilbert curve (H).
- 25 6. The method according to claim 5, wherein said rot transformation indicates
cyclically shifting the bits of said bitblock.
- 30 7. The method of claim 5 or 6, wherein organizing is at least one of searching,
sorting, storing, retrieving, inserting, deleting, querying, range querying, data
elements in said database system.
- 35 8. A computer-readable data storage medium for storing program code for
executing, when being loaded into a computer, the method according to one of
claims 5 to 7.

x=	0	1	2	3	4	5	6	7
y= 0	0	1	4	5	16	17	20	21
y= 1	2	3	6	7	18	19	22	23
y= 2	8	9	(12	13)	24	25	28	29
y= 3	10	11	(14	15)	26	27	30	31
y= 4	32	33	(36	37)	48	49	52	53
y= 5	34	35	(38	39)	50	51	54	55
y= 6	40	41	44	45	56	57	60	61
y= 7	42	43	46	47	58	59	62	63

Fig. 1

x=	0	1	2	3	4	5	6	7
y=0	0	1	14	15	16	19	20	21
y=1	3	2	13	12	17	18	23	22
y=2	4	7	8	11	30	29	24	25
y=3	5	6	9	10	31	28	27	26
y=4	58	57	54	53	32	35	36	37
y=5	59	56	55	52	33	34	39	38
y=6	60	61	50	51	46	45	40	41
y=7	63	62	49	48	47	44	43	42

Fig. 2

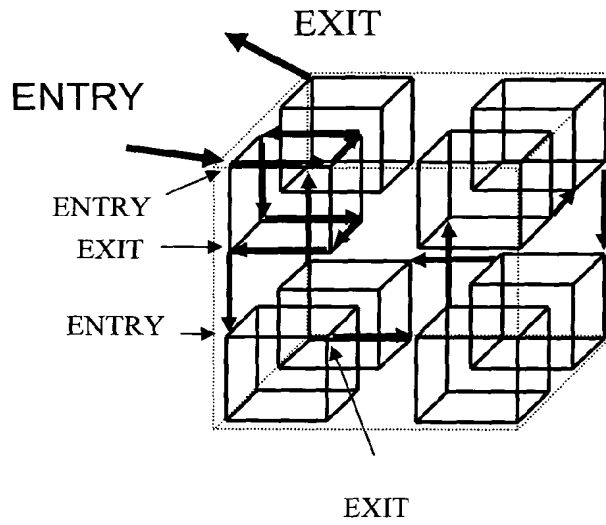


Fig. 3

Expl.1 Expl.2: Expl.3: Expl.4: Expl.5:

f: 101

0: 00	0: 000	0: 101	0: 011	0: 100
1: 01	1: 001	1: 100	1: 001	1: 110
2: 11	2: 011	2: 110	2: 101	2: 111
3: 10	3: 010	3: 111	3: 111	3: 011
	4: 110	4: 011	4: 110	4: 001
	5: 111	5: 010	5: 100	5: 000
	6: 101	6: 000	6: 000	6: 011
	7: 100	7: 001	7: 010	7: 101

Fig. 4

```

zyx      zyx  Tab(0)=(000/+2)
0 | 000(a) | 000(a) 000 (a)
000      ...  010
000      ...  110
000      ...  100
000      ... >> 101
000      ...  111
000      ...  011
000      ..1(b) 001
-----
1 001      ..0(b)
001      ...
...      ...
001      ..1(b)
-----
2 011      ..0(b)
...      ...
011      ..0(b)
-----
3 010      ..1(b) 011 Tab(3)=(011/0)
...      ...
...      >>      000 XOR 011=011
...      >>      100 XOR 011=111
010      1..(b) 111
-----
4 110      0..(b)
...      ...
110      ..1(b)
-----
5 111      ..0(b)
...      ...
111      ..0(b)
-----
6          Tab(6)=(110/+1)
101      ..1(b) 0 | 110(a) 110(a)
          ...
          ...
101          1 010
101          2 011
          Sub-Sub-Cube:
          ConcatTab(3)=(011/0)with
          Tab(6)=(110/+1)
101          3 111 011
          3 111 010
          3 111 000
          3 111 001
          3 111 101
          3 111 100
          3 111 110
          3 111 111
101          4 101
101          5 001
101          6 000
101      ..0(b) 7 | 100(a) 100 (a)
-----
7 100      ..1(b)
...      ...
| 100      | 100(a)

```

Fig. 5

Primitive 1D, 2 bit, Data Cube:

1st bit: 0----->1

2nd bit: 0-->1-->0-->1

z	z	yz	yz	yz	yz	xyz	xyz
0	0	00	00	00	00	000	000
	1		01		01		001
1	0	01	00	01	00	001	000
	1		>01<		>10<		010

		11	>11<	11	>00<	011	000
			10		10		010
		10	11	10	11	010	011
			10		10		>111<
=====							
						110	>011<
							111
						111	110
							100

						101	110
							100
						100	101
							100

(a)

(b)

(c)

(d)

Fig. 6

Dr.-Ing. Hermann TROPF
February 18, 2003
T50310EP

EPO - Munich
41
18. Feb. 2003

Abstract

A database system and method is provided for organizing data elements according to a Hilbert curve, said data elements being representable by a plurality of components, said database system comprising:

first means for generating a plurality of bitblocks by bitwise interleaving the components of the data elements;

second means for applying a fliprot transformation to a first bitblock;

said fliprot transformation comprising a flip transformation and a rot transformation, said flip transformation indicating which bits of said bitblock are to be inverted, said rot transformation indicating which bits of said bitblock are to be interchanged;

third means for obtaining, for each further bitblock, a fliprot transformation by a concatenation of two or more fliprot transformations; and

fourth means for applying each fliprot transformation to its corresponding bitblock;

whereby the bitblock bits determine the organization of said data elements according to said Hilbert curve.

(Fig. 5)